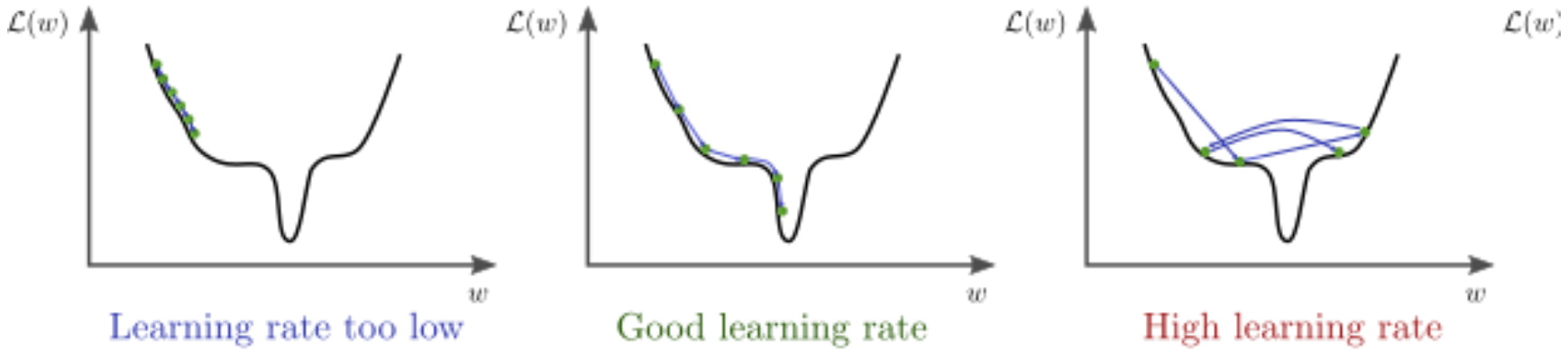
Learning Rate Decay

As we saw, based on the learning rate we can have the following situation:



Learning Rate Decay

$$\alpha = \frac{1}{1 + decay - rate * epoch - num}\alpha_0$$

Example:

$\alpha_0 = 0.2; decay - rate = 1$

Epoch	α
1	0.1
2	0.067
3	0.05
4	0.04
…	…

Other formulas:

$\alpha = \beta^{epoch-num}\alpha_0$ with $\beta \in (0,1)$ - Exponentially Decay

$\alpha = \frac{k}{\sqrt{epoch-num}}\alpha_0$ with $k > 1$